Network Working Group
Request for Comments: 1133

J. Yu
H-W. Braun
Merit Computer Network
November 1989

Routing between the NSFNET and the DDN

Status of this Memo

This document is a case study of the implementation of routing between the NSFNET and the DDN components (the MILNET and the ARPANET). We hope that it can be used to expand towards interconnection of other Administrative Domains. We would welcome discussion and suggestions about the methods employed for the interconnections. No standards are specified in this memo. Distribution of this memo is unlimited.

1. Definitions for this document

The NSFNET is the backbone network of the National Science Foundation's computer network infrastructure. It interconnects multiple autonomously administered mid-level networks, which in turn connect autonomously administered networks of campuses and research centers. The NSFNET connects to multiple peer networks consisting of national network infrastructures of other federal agencies. One of these peer networks is the Defense Data Network (DDN) which, for the sake of this discussion, should be viewed as the combination of the DoD's MILNET and ARPANET component networks, both of which are national in scope.

It should be pointed out that network announcements in one direction result in traffic the other direction, e.g., a network announcement via a specific interconnection between the NSFNET to the DDN results in packet traffic via the same interconnection between the DDN to the NSFNET.

2. NSFNET/DDN routing until mid '89

Until mid-1989, the NSFNET and the DDN were connected via a few intermediate routers which in turn were connected to the ARPANET. These routers exchanged network reachability information via the Exterior Gateway Protocol (EGP) with the NSFNET nodes as well as with the DDN Mailbridges. In the context of network routing these Mailbridges can be viewed as route servers, which exchange external network reachability information via EGP while using a proprietary protocol to exchange routing information among themselves. Currently, there are three Mailbridges at east coast locations and

three Mailbridges at west coast locations. Besides functioning as route servers the Mailbridges also provide for connectivity, i.e, packet switching, between the ARPANET and the MILNET.

The intermediate systems between the NSFNET and the ARPANET were under separate administrative control, typically by a NSFNET mid-level network.

For a period of time, the traffic between the NSFNET and the DDN was carried by three ARPANET gateways. These ARPANET gateways were under the administrative control of a NSFNET mid-level network or local site and had direct connections to both a NSFNET NSS and an ARPANET PSN. These routers had simultaneous EGP sessions with a NSFNET NSS as well as a DDN Mailbridge. This resulted in making them function as packet switches between the two peer networks. As network routes were established packets were switched between the NSFNET and the DDN.

The NSFNET used three NSFNET/ARPANET gateways which had been provided by three different sites for redundancy purposes. Those three sites were initially at Cornell University, the University of Illinois (UC), and Merit. When the ARPANET connections at Cornell University and the University of Illinois (UC) were terminated, a similar setup was introduced at the Pittsburgh Supercomputer Center and at the John von Neumann Supercomputer Center which, together with the Merit connection, allowed for continued redundancy.

As described in RFC1092 and RFC1093, NSFNET routing is controlled by a distributed policy routing database that controls the acceptance and distribution of routing information. This control also extends to the NSFNET/ARPANET gateways.

2.1 Inbound announcement -- Routes announced from the DDN to the NSFNET

In the case of the three NSFNET/ARPANET gateways, each of the associated NSSs accepted the DDN routes at a different metric. The route with the lowest metric then was favored for the traffic towards the specific DDN network, but had that specific gateway to the DDN experienced problems with loss of routing information, one of the redundant gateways would take over and carry the load as a fallback path. Assuming consistent DDN routing information at any of the three gateways, as received from the Mailbridges, only a single NSFNET/ARPANET gateway was used at a given time for traffic from the NSFNET towards the DDN, with two further gateways standing by as hot backups. The metric for network announcements from the DDN to the NSFNET was coordinated by the Merit/NSFNET project.

2.2 Outbound announcement -- Routes announced from the NSFNET to the DDN

Each NSS involved with NSFNET/DDN routing had an EGP peer relation with the NSFNET/ARPANET gateway. Via EGP it announced a certain set of NSFNET connected networks, again, as controlled by the distributed policy routing database, to its peer. The NSFNET/ARPANET gateway then redistributed the networks it had learned from the NSS to the DDN via a separate EGP session. Each of the NSFNET/ARPANET gateways used a separate Autonomous System number to communicate EGP information with the DDN. Also these Autonomous System numbers were not the same as the NSFNET backbone uses to communicate with directly attached client networks. The NSFNET/ARPANET gateways used the Autonomous System number of the local network. The metrics for announcing network numbers to the DDN Mailbridges were set according to the requests of the mid-level network of which the specific individual network was a client. Mid-level network also influenced the specific NSFNET/ARPANET gateway used, including primary/secondary selection. These primary/secondary selections among the NSFNET/ARPANET gateways allowed for redundancy, while the preference of network announcements was modulated by the metric used for the announcements to the DDN from the NSFNET/ARPANET gateways. Some of the selection decisions were based on reliability of a specific gateway or congestion expected in a specific PSN that connected to the NSFNET/ARPANET gateway.

2.3 Administrative aspects

From an administrative point of view, the NSFNET/ARPANET gateways were administered by the institution to which the gateway belonged. This has never been a real problem due to the excellent cooperation received from all the involved sites.

3. NSFNET/DDN routing via attached Mailbridges

During the first half of 1989 a new means of interconnectivity between the NSFNET and the DDN was designed and implemented. Ethernet adapters were installed in two of the Mailbridges, which previously just connected the MILNET and the ARPANET, allowing a direct interface to NSFNET nodes. Of these two Mailbridges one is located on the west coast at NASA-Ames located at Moffett Field, CA, and the other one is located on the east coast at Mitre in Reston, VA. With this direct interconnection it became possible for the NSFNET to exchange routing information directly with the DDN route servers, without a gateway operated by a mid-level network in the middle. This also eliminated the need to traverse the ARPANET in order to reach MILNET sites. It furthermore allows the Defense Communication Agency as well as the National Science Foundation to

exercise control over the interconnection on a need basis, e.g., the connectivity can now be easily disabled from either site at times of tighter network security concerns.

3.1 Inbound announcement -- Routes announced from the DDN to the NSFNET

The routing setup for the direct Mailbridge connections is somewhat different, as compared to the previously used NSFNET/ARPANET gateways. Instead of a single NSFNET/ARPANET gateway carrying all the traffic from the DDN to the NSFNET at any moment, the distribution of network numbers is now split between the two Mailbridges. This results in a distributed load, with specific network numbers always preferring a particular Mailbridge under normal operating circumstances. In the case of an outage at one of the Mailbridge connections, the other one fully takes over the load for all the involved network numbers. For this setup, the two DDN links are known as two different Autonomous System numbers by the NSFNET. The routes learned via the NASA-Ames Mailbridges are part of the Autonomous System 164 which is also the Autonomous System number which the Mailbridges are using by themselves during the EGP session. In the case of the EGP sessions with the Mitre Mailbridge, the DDN-internal Autonomous System number of 164 is overwritten with a different Autonomous System number (in this case 184) and the routes learned via the Mitre Mailbridge will therefore become part of Autonomous System 184 within the NSFNET.

The NSFNET-inbound routing is controlled by the distributed policy routing database. In particular, the network number is verified against a list of legitimate networks, and a metric is associated with an authorized network number for a particular site. For example, both NSSs in Palo Alto and College Park learn net 10 (the ARPANET network number) from the Mailbridges they are connected to and have EGP sessions. The Palo Alto NSS will accept Net 10 with a metric of 10, while the College Park NSS will accept the same network number with a metric of 12. Therefore, traffic destinated to net 10 will prefer the path via the Palo Alto NSS and the NASA-Ames Mailbridge. If the connection via the NASA-Ames Mailbridge is not functioning, the traffic will be re-routed via the Mailbridge link at Mitre. Each of the two NSS accepts half of the network routes via EGP from its co- located Mailbridge at a lower metric and the other half at a higher metric. The half with the lower metric at the Palo Alto NSS will be the same set which uses a higher metric at the College Park NSS and vice versa.

There are at least three different possibilities about how the NSFNET could select a path to a DDN network via a specific Mailbridge, i.e., the one at NASA-Ames versus the one at Mitre:

1. Assign a primary path for all DDN networks to a single Mailbridge and use the other purely as a backup path.

2. Distribute the DDN networks randomly across the two Mailbridges.

3. Let the DDN administration inform the NSFNET which networks on the DDN are closer to a specific Mailbridge so that the particular Mailbridge would accept these networks at a lower metric. The second Mailbridge would then function as a backup path. From a NSFNET point of view, this would mean treating the DDN like other NSFNET peer networks such as the NASA Science network (NSN) or DOE's Energy Science Network (ESNET).

We are currently using alternative (2) as an interim solution. At this time, the DDN administration is having discussions with NSFNET about moving to alternative (3), which would allow them control over how the DDN networks would be treated in the NSFNET.

3.2 Outbound announcement -- Routes announced from the NSFNET to the DDN

The selection of metrics for announcements of NSFNET networks to the DDN is controlled by the NSFNET. The criteria for the metric decisions is based on distances between the NSS, which introduces a specific network into the NSFNET, and either one of the NSSs that has a co-located Mailbridge. In this context, the distance translates into the hop count between NSSs in the NSFNET backbone. For example, the Princeton NSS is currently one hop away from the NSS co-located with the Mitre Mailbridge, but is three hops away from the NSS with the NASA-Ames Mailbridge. Therefore, in the case of networks with primary paths via the Princeton NSS, the Mitre Mailbridge will receive the announcements for those networks at a lower metric than the NASA-Ames Mailbridge. This means that the traffic from the DDN to networks connected to the Princeton NSS will be routed through the Mailbridge at Mitre to the College Park NSS and then through the Princeton NSS to its final destination. This will guarantee that traffic entering the NSFNET from the DDN will take the shortest path to its NSFNET destination under normal operating conditions.

3.3 Administrative aspects

Any of the networks connected via the NSFNET can be provided with the connectivity to the DDN via the NSFNET upon request from the mid-level network through which the specific network is connected.

For networks that do not have a DDN connection other than via NSFNET, the NSFNET will announce the nets via one of the Mailbridges with a

low metric to create a primary path (e.g., metric "1") and via the second Mailbridge as a secondary path (e.g., metric "3"). For networks that have their own DDN connection and wish to use the NSFNET as a backup connection to DDN, the NSFNET will announce those networks via the two Mailbridges at higher metrics.

The mid-level networks need to make a specific request if they want client networks to be announced to the DDN via the NSFNET. Those requests need to state whether this would be a primary connection for the specific networks. If the request is for a fallback connection, it needs to state the existing metrics in use for announcements of the network to the DDN.

4. Shortcomings of the current NSFNET/DDN interconnection routing

The current setup makes full use of the two Mailbridges that connect to the NSFNET directly, with regard to redundancy and load sharing. However, with regard to performance optimization, such as packet propagation delay between source/destination pairs located on disjoint peer networks, there are some shortcomings. These shortcomings are not easy to overcome because of the limitations of the current architecture. However, it is a worthwhile topic for discussion to aid future improvements.

To make the discussion easier, the following assumptions and terminology will be used:

The NSFNET is viewed as a cloud and so is the DDN. The two have two connections, one at east coast and one at west coast.

mb-east -- the Mailbridge at Mitre

mb-west -- the Mailbridge at Ames

NSS-east -- the NSS egp peer with mb-east

NSS-west -- the NSS egp peer with mb-west

DDN.east-net -- networks connected to DDN and physically closer to mb-east

DDN.west-net -- networks connected to DDN and physically closer to mb-west

NSF.east-net -- networks connected to NSFNET and physically closer to NSS-east

NSF.west-net -- networks connected to NSFNET and physically closer

to NSS-west

The traffic between NSFNET<->DDN will fall into the following patterns:

a) NSF.east-net <-> DDN.east-net or
   NSF.west-net <-> DDN.west-net

b) NSF.east-net <-> DDN.west-net or
   NSF.west-net <-> DDN.east-net

The ideal traffic path for a) and b) should be as follows:

For traffic pattern a)

NSF.east-net<-->NSS.east<-->mb-east<-->DDN.east-net

or

NSF.west-net<-->NSS.west<-->mb-west<-->DDN.west-net

For traffic pattern b)

```
NSF.east-net-*->NSS.west-->mb-west-->DDN.west-net-**->mb-east
                                                      |
                                                NSF.east-net<--NSS-east
```

or

```
NSF.west-net-*->NSS.east-->mb-east-->DDN.east-net-**->mb-west
                                                      |
                                                NSF.west-net<--NSS-west
```

Note:

-*-> is used to indicate traffic transcontinentally traversing the NSFNET backbone

-**-> is used to indicate traffic transcontinentally traversing the DDN backbone

The traffic for a) will transcontinentally traverse NEITHER the NSFNET backbone, NOR the DDN backbone.

The traffic for b) will transcontinentally traverse NSFNET once and DDN once and only once for each.

For the current set up,

The traffic path for pattern a) would have chances to transcontinentally traverse both NSFNET and DDN.

The traffic path for pattern b) would have chances to transcontinentally traverse the DDN in both directions.

To achieve the ideal traffic path it requires the NSFNET to implement (3) as stated above, i.e., to treat the DDN like other NSFNET peer or mid-level networks. As mentioned before, discussions between NSFNET and DDN people are underway and the DDN is considering providing the NSFNET with the required information to accomplish the outlined goals in the near future.

At such time as this is accomplished, it will reduce the likelihood of packet traffic unnecessarily traversing national backbones.

One of the best ways to optimize the traffic between two peer networks, not necessary limited to the NSFNET and the DDN, is to try to avoid letting traffic traverse a backbone with a comparatively slower speed and/or a backbone with a significantly larger diameter network. For example, in the case of traffic between the NSFNET and the DDN, the NSFNET has a T1 backbone and a maximum diameter of three hops, while the DDN is a relatively slow network running largely at 56Kbps. In this case the overall performance would be better if traffic would traverse the DDN as little as possible, i.e., whenever the traffic has to reach a destination network outside of the DDN, it should find the closest Mailbridge to exit the DDN.

The current architecture employed for DDN routing is not able to accomplish this. Firstly, the technology is designed based on a core model. It does not expect a single network to be announced by multiple places. An example for multiple announcements could be two NSSs announcing a single network number to the two Mailbridges at their different locations. Secondly, the way all the existing Mailbridges exchange routing information among themselves is done via a protocol similar to EGP, and the information is then distributed via EGP to the DDN-external networks. In this case the physical distance information and locations of network numbers is lost and neither the Mailbridges nor the external gateways will be able to do path optimization based on physical distance and/or propagation delay. This is not easy to change, as in the DDN the link level routing information is decoupled from the IP level routing, i.e., the IP level routing has no information about topology of the physical infrastructure. Thus, even if an external gateway to a DDN network were to learn a particular network route from two Mailbridges, it would not be able to favor one over the other DDN exit point based on

the distance to the respective Mailbridge.

5. Conclusions

While recent changes in the interconnection architecture between the NSFNET and DDN peer networks have resulted in significant performance and reliability improvements, there are still possibilities for further improvements and rationalization of this inter-peer network routing. However, to accomplish this it would most likely require significant architectural changes within the DDN.

6. References

[1] Rekhter, Y., "EGP and Policy Based Routing in the New NSFNET Backbone", RFC 1092, IBM Research, February 1989.

[2] Braun, H-W., "The NSFNET Routing Architecture", RFC 1093, Merit/NSFNET Project, February 1989.

[3] Collins, M., and R. Nitzan, "ESNET Routing", DRAFT Version 1.0, LLNL, May 1989.

[4] Braun, H-W., "Models of Policy Based Routing," RFC 1104, Merit/NSFNET Project, February 1989.

Security Considerations

Security issues are not addressed in this memo.

Authors' Addresses

Jessica (Jie Yun) Yu
Merit Computer Network
1075 Beal Avenue
Ann Arbor, Michigan 48109

Telephone: 313 936-2655
Fax: 313 747-3745
EMail: jyy@merit.edu

Hans-Werner Braun
Merit Computer Network
1075 Beal Avenue
Ann Arbor, Michigan 48109

Telephone: 313 763-4897
Fax: 313 747-3745
EMail: hwb@merit.edu